

Poster III-46

Formalized Representation of Scientific Literature Using Biological Association Networks

***Yuryev, Anton, Daraselia, Nikolai, Nikitin, Alexander, Egorov, Sergei, Mazo, Ilya
Ariadne Genomics Inc., Rockville, MD, USA***

The major part of biological knowledge is stored in the form of scientific literature. The algorithmic analysis of vast biological pathways can only be done if data is formatted or in a database. Various software tools and algorithms analyzing properties of individual proteins, protein classes or global characteristics of biological networks could access such database.

The Natural Language Processing (NLP) was used to convert knowledge from scientific texts into the relational database (ResNet). Currently, our NLP extracts facts about biomolecular interactions, gene expression regulation, molecular biosynthesis and trafficking and cell process regulation.

ResNet database stores close to 100,000 facts about 14,251 proteins, protein complexes, cell organelles, compartments and processes. The database can be accessed using PathwayAssist client software. PathwayAssist allows visualization of biological pathways in graphical form using both automated and manual layouts. The client has algorithms to build and expand biological pathways and to allow analysis and data-mining in PathwayAssist.

We are using ResNet data to develop algorithm for analysis of Gene Expression data, statistical analysis of biological pathways, and integration of regulatory, physical interaction and biochemical cellular networks.